COMMUNITY ECOLOGY 8(2): 239-246, 2007
1585-8553/$20.00 © Akadémiai Kiadó, Budapest
DOI: 10.1556/ComEc.8.2007.2.10

Geographic and taxonomic bias in land snail distribution data of Hungary

P. Sólymos

Department of Ecology, Faculty of Veterinary Science, Szent István University, Rottenbiller st. 50, H-1077 Budapest, Hungary. Tel.: +36 1 4784254, Fax: +36 1 4784232, E-mail: Solymos.Peter@aotk.szie.hu

Keywords: Biodiversity, Biotic impediment, Gastropoda, Mollusca, Protected areas, Spatial autocorrelation, Spatial coverage, Spatial intensity.

Abstract: The importance of accurate species databases is debated in the recent literature of biodiversity assessment, considering that limited resources for conservation could be better allocated to assessment based on cost effective biodiversity features. I aimed to provide an understanding of sampling bias and provide practical advice to minimize bias either before or after data collection. I used 10×10 km^2 UTM grid data for 121 land snail species to account for geographic and taxonomic sampling bias in Hungary. Sampling intensity corrected for species richness varied significantly among regions, although regions were not good predictors of sampling intensity. Residuals were significantly autocorrelated in 15 km distance, indicating small scale heterogeneity in sampling intensity compared to species richness. Sampling coverage and intensity were higher close to human settlements and sampling intensity was higher within protected areas than outside. Commonness of species was positively associated with sampling intensity, while some rare species were over-represented in the records. Sampling intensity of microsnails (<3 mm) was significantly lower than that of the more detectable large species (>15 mm). Systematic effects of the collecting methods used in malacological research may be responsible for these differences. Understanding causes of sampling bias may help to reduce its effects in ecological, biogeographical and conservation biological applications, and help to guide future research.

Nomenclature: Bank (2004).

Introduction

The preservation of biological diversity needs good quality data and effective methods (Williams et al. 2002). Although invertebrates make up a high proportion of total species richness, most invertebrate taxa are generally under-represented in biodiversity assessment and conservation planning compared to well-studied vertebrates and butterflies. Under-representation of invertebrates is due to the lack or low quality of distribution data (Pressey 2004).

There is a debate on the need for accurate and comprehensive species databases. Brooks et al. (2004) argue that species are the only valid currency for conservation assessment, while according to Cowling et al. (2004) the appropriate biodiversity features (environmental surrogates, land classes) to be used in conservation assessment depend on the goal, spatial scale, implementation opportunities and level of biodiversity knowledge in a particular region. The distribution database of land snails is the most comprehensive invertebrate database in Hungary, and it thus provides a basis for studying the usefulness of the data in biodiversity assessment.

The utility of the distribution data depends heavily on the representativeness of the data – as a sample – to the entire sampling universe, the biota of a given geographic area. In the real world, distribution data bases are not always as representative to the whole as desired, and have inherent bias due to geographic and taxomical variation in sampling effort referred to Wallacean and Linnean shortfalls, respectively (Whittaker et al. 2005). Wallacean shortfall (Lomolino 2004) means that many areas are seriously under-collected for most taxa, even vascular plants and birds, so that biased datasets may confound biogeographic patterns (Nelson et al. 1990) or the results of area selection (Reddy and Dávalos 2003). Linnean shortfall (Brown and Lomolino 1998) means that much of the diversity we do know about has yet to be formally described and catalogued. In a wider sense, biased faunal and floral lists – because of varying rarity and detection probability of the species – are also special cases of Lin-

nean shortfalls and can affect biogeographic and conservation applications of the data.

These two types of bias are widespread in large scale studies (e.g., Ramsey and Shultz 2004) and apparent in biotic inventories (e.g., Stohlgren et al. 1995). Bias may have a strong influence on site selection applications (Freitag and Van Jaarsveld 1998), yet relatively little work has been done so far in explicitly considering the causes and consequences of bias (Whittaker et al. 2005). Reddy and Dávalos (2003) analysed the effects of geographic sampling bias, while taxonomic bias has not been evaluated in such a framework on a large (e.g., national) scale.

I used the distribution of Hungarian land snails to test for geographic and taxonomic bias. I analysed regional variation in sampling intensity, and the effect of the spatial arrangement of human settlements and protected areas on sampling intensity. I tested the effect of rarity and detection probability of land snail species on their sampling intensity. I aimed to provide an understanding of sampling bias and practical advice to minimize them either before or after data collection.

Materials and methods

The area of Hungary is covered by 1052 10 × 10 km^2 UTM grid cells. Smaller cells occurred near the country's borders or in the spherical correction zone of the UTM system. I grouped the cells into regions according to Dévai and Miskolczi (1987): (1) Tisza Plain, (2) Danube Plain, (3) Lesser Hungarian Plain, (4) Western Marginal Area, (5) Transdanubian Hills, (6) Transdanubian Mountains and (7) Northern Mountains (Fig. 1).

Distribution data of land snails were derived from Pintér et al. (1979), Pintér and Szigethy (1979, 1980) and Fehér and Gubányi (2001). Invasive and introduced species were excluded because their conservation relevance is doubtful (Patten and Erickson 2001). Slugs were also excluded due to collecting and identification problems (Cameron and Pokryszko 2005). In total, 121 species were involved in the analysis (including semislugs).

Distribution data were available for 704 (66.9%) of the 1052 UTM cells. The Mollusca Collection of the Hungarian Natural History Museum (Fehér and Gubányi 2001) contained 26023 lots (a lot being a collection of one species from one location collected on one occasion) from 612 (58.2%) UTM cells.

I estimated spatial sampling intensity as the residuals of the logarithm of number of museum lots per UTM cell regressed against logarithm of total species richness within that cell. I used general linear model to assess regional differences in spatial sampling intensity. Based on spatial coordinates of the centroids of the UTM cells, I tested the linear model residuals for spatial autocorrelation using Moran's *I* and 15 km spatial lag to include centroids of all eight potential neighbours of the UTM cells. Out of the 612 UTM cells with intensity data, I used only 606 cells that had at least one neighbour cell. I made spatial autocorrelograms for the number of species, number of museum lots, spatial sampling intensity (as defined

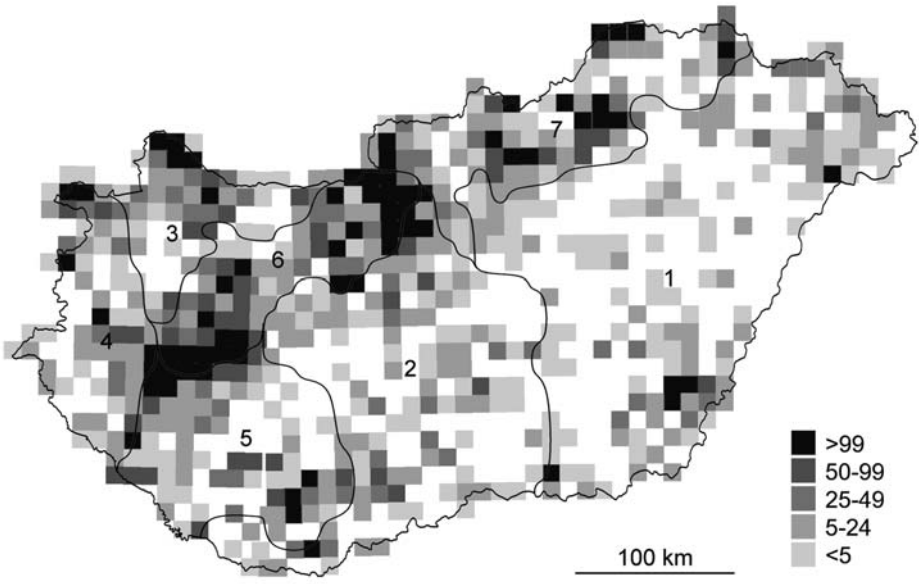

Figure 1. Number of museum lots in 10 × 10 km^2 UTM grid cells in Hungary based on the land snail collection of the Hungarian Natural History Museum (Fehér and Gubányi 2001). Shades of grey represent levels of sampling intensity (number of museum lots per UTM cell) as indicated in the figure, white areas are unexplored. Numbers represent major geographic units: (1) Tisza Plain, (2) Danube Plain, (3) Lesser Plain, (4) Western Marginal Area, (5) Transdanubian Hills, (6) Transdanubian Mountains and (7) Northern Mountains.

above) and linear model residuals based on 15 km spatial lags. For spatial statistics, I used the *spdep* package (Bivand 2007) of the R software (R Development Core Team 2007).

I analysed the effect of human settlements on the pattern of sampling following Reddy and Dávalos (2003) by measuring the distance between a given UTM grid cell containing distribution data and the UTM cell containing the nearest town centre. I identified UTM cells containing the centres of 252 towns in Hungary based on the data of Miskolczi et al. (1997) and the Hungarian Central Statistical Office (KSH) (2001). I compared the shape of the distribution and central tendency of the distances to the nearest city for the 704 UTM grid cells containing distribution data, and for 704 cells identified randomly without replacement. Besides spatial coverage (distribution of the 704 cells with data), I compared the observed intensity data (26023 data items) with the same number of data gained from randomization, similarly as for coverage but with replacement. To compare distributions, I used the Kolmogorov-Smirnov two-sample test and to compare central tendency I used the Mann-Whitney *U*-test and approximate randomisation test because of the large number of cases (Potvin and Roff 1993).

I followed Reddy and Dávalos (2003) to test the effect of the distribution of protected areas on the sampling intensity of the species. Sampling intensity for a species was the mean number of lots of that species per cell in which it was present. I compared sampling intensity inside and outside UTM cells containing protected areas. Because both variables are subject to error, I used Type II regression (standard major axis regression, Sokal and Rohlf 1995). If the sampling intensity inside and outside the protected areas was identical, the slope of the regression would not differ significantly from unity. I computed 95% confidence intervals to test the departure of the slope parameter beta from one. The effect of area was controlled both inside and outside, by using sampling intensity per unit area.

Sampling intensity of the species per unit area followed a lognormal distribution (Shapiro-Wilk test, $W = 0.985$, $n = 116$, $p = 0.216$), so for the analysis I used log-transformed values of sampling intensity. I tested the effect of rarity (frequency of occupied UTM cells) and detection probability (shell size) with a general linear model. I assessed shell size as the greatest value of shell height or width obtained from Kerney et al. (1983) ($n = 97$, semislugs were excluded because their shell size is not related to their body size).

I categorised the species into five classes according to their rarity and shell size independently, with (almost) equal numbers of cases in each class. Then, I used these categories as treatment levels of fixed factors in the model, also testing their interactions. Log-transformed sampling intensity of the species per unit area was the response variable. I used the least significant differences test for pairwise comparisons, and significance levels were corrected for multiple comparisons by the false discovery rate (FDR) method (Benjamini et al. 2001).

Results

The distribution of the number of museum lots within UTM cells was right-skewed (skewness = 6.73 ± 0.099 SE, $n = 612$). Only 69 cells out of the 612 (6.6%) were represented by 100 or more museum lots (Fig. 1). The number of museum lots per cell showed peaks in the Northern Mountains, in some sporadically distributed cells in the Danube and Tisza Plains (the areas around of the Bátorliget Nature Reserve, the towns of Békéscsaba and Szeged), in the surroundings of Budapest, the Bakony Mountains and the town of Keszthely, in some areas near the western border (Szigetköz, Soproni Mts and Kőszegi Mts), and in areas around Pécs in southern Transdanubia (Fig. 1).

I found significant variation in sampling intensity (residuals of number of museum lots after correcting for species richness) among regions ($F = 2.12$, df = 7, 599, $p < 0.05$, $R^2 = 0.024$). Sampling intensity was significantly ($p < 0.01$) higher in the Lesser Hungarian Plains than in other regions. Negative peaks were not analysed, because empty cells were not included. According to sampling coverage, empty cells were most frequent in the Tisza and Duna Plains (Fig. 1).

The low R^2 value of the linear model indicated that regions could not account for most of the variation in sampling intensity. The autocorrelation in the regression residuals was significantly greater than expected under the null model of no autocorrelation (global Moran's $I = 0.152$, expectation under null model = -0.009, $z = 6.07$, $p < 0.001$).

Spatial autocorrelation of the number of species and the number of museum lots per UTM cell was significant within 3-4 spatial lags (45-60 km distance). The autocorrelation of richness corrected sampling intensity and the residuals of the linear model was significant within one lag (15 km distance) (Fig. 2).

The effect of the location of towns on sampling coverage (cells with data) and intensity was significant. The average distance between UTM cells containing distribu-

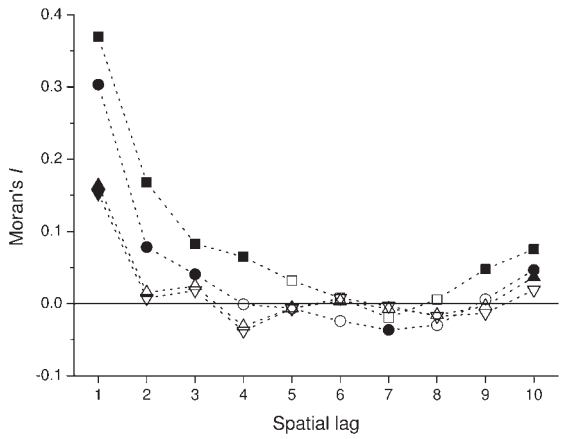

Figure 2. Moran's *I* spatial correlogram of (■) number of species, (●) number of museum lots, (▲)sampling intensity corrected for species richness and (▼) residuals of the general linear model (sampling intensity vs. regions) within 10 × 10 km^2 UTM cells. One spatial lag represents 15 km distance, full symbols indicate significant ($p < 0.05$), open symbols indicate non-significant Moran's *I* values.

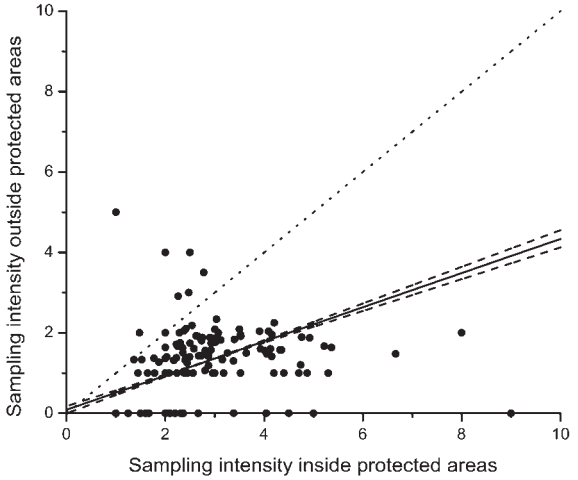

Figure 4. Sampling intensity of the testaceous species (n = 97) inside and outside 10 × 10 km^2 UTM grid cells containing protected areas. The line indicates major axis of Type II regression ($y = 0.505x + 0.128$). Broken lines correspond to 95% confidence limits at this axis, dotted line indicates the slope of unity indicating that sampling intensity inside and outside protected areas are equal.

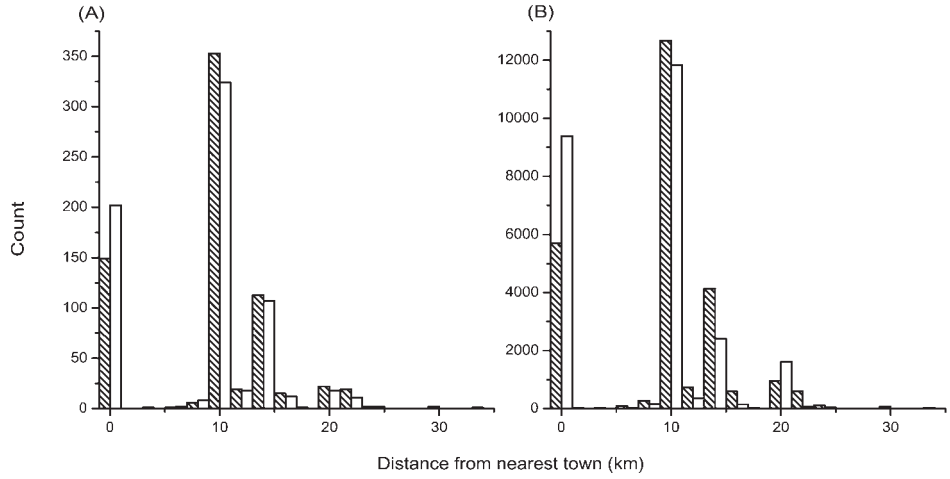

Figure 3. Distribution of the distances to the nearest town for (**A**) sampling coverage (distribution of n = 704 UTM cells containing distribution data) and (**B**) sampling intensity (n = 26023) data at 10 × 10 km^2 resolution. Empty columns correspond to observed data, hatched columns correspond to data from random permutations (**A**) without and (**B**) with replacement.

tion data and the nearest towns (8.07 ± 5.745 km SD, n = 704) was significantly smaller than the average distance of the randomised data set without replacement (9.14 ± 5.705 km SD, n = 704) (Mann-Whitney test, U = 225654.5, $p < 0.01$; Kolmogorov-Smirnov two-sample test, $D = 0.078$, $p < 0.05$; Fig. 3a). The average distance between UTM cells of each museum lot and the nearest city (7.19 ± 6.005 km SD, n = 26023) was significantly smaller than the average distance of the random data set with replacement (9.07 ± 5.863 km SD, n = 26023) (approximate randomisation test with 5000 permutations, $p < 0.001$; Kolmogorov-Smirnov two-sample test, $D = 0.166$, $p < 0.001$; Fig. 3b).

Sampling intensity of the species was higher inside protected areas than outside. The estimated slope parameter of the major axis of the Type II regression was significantly lower than one ($\beta = 0.425$, 95% $CL_1 = 0.395$, 95% $CL_2 = 0.456$, $y = 0.425\ x + 0{,}088$, $R^2 = 0{,}0041$, n = 97, including only testaceous species). Some species (*Platyla polita*, *Truncatellina callicratis*, *Macrogastra plicatula*, *Bulgarica vetusta*, *Urticicola umbrosus*, *Kovacsia kovacsi*, *Helix lutescens*) showed the opposite trend from

Table 1. Effects of rarity (area of occupancy) and shell size on the sampling intensity of the species[a].

Source of variation	df	SS	MS	*F*
With outliers				
Between groups	25	20.879	0.835	46.72***
Rarity	4	0.155	0.039	2.17+
Shell size	4	0.288	0.072	4.02**
Interaction	16	0.608	0.038	2.12*
Error	91	1.627	0.018	
Total	116	22.505		
Without outliers				
Between groups	25	18.833	0.753	60.252***
Rarity	4	0.348	0.009	6.955***
Shell size	4	0.236	0.006	4.722**
Interaction	16	0.303	0.002	1.514ns
Error	86	1.075	0.001	
Total	111	19.908		

[a]Two-way general linear model, sampling intensity was log-transformed, ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, +: $p < 0.1$, ns: not significant.

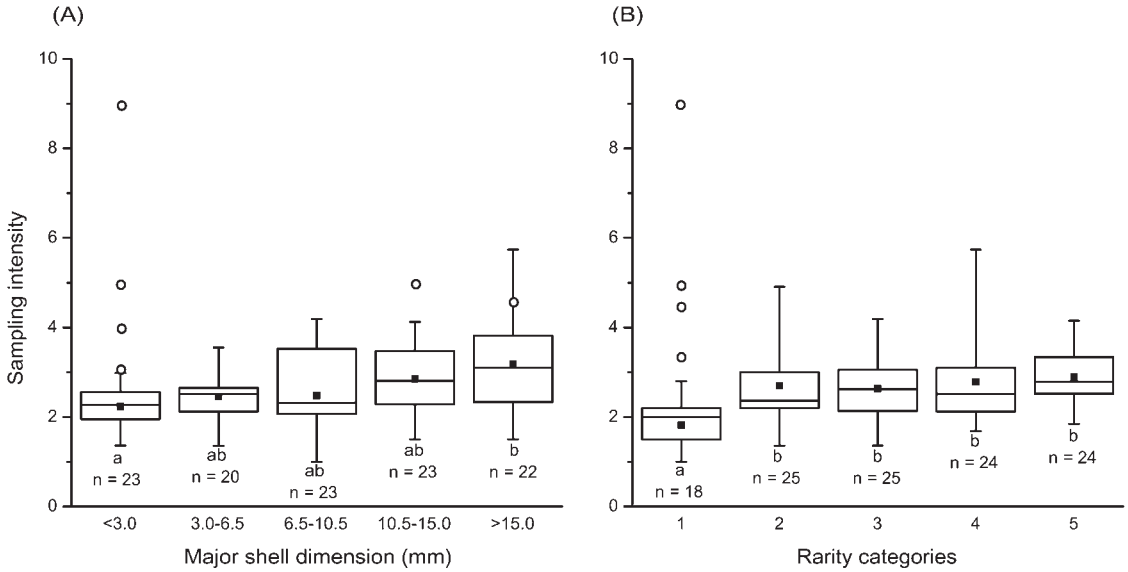

Figure 5. Effect of (**A**) shell size and (**B**) rarity on the sampling intensity of species (full square: mean, central line: median, box: inter quartile range, whiskers: minimum and maximum, circles: outliers; outliers were excluded from the calculation of descriptive statistics). Letters indicate significant ($p < 0.05$) differences (least significant differences, significance levels were corrected by the FDR method).

the majority of species, with sampling intensity of these species higher outside protected areas (Fig. 4).

Rarity and shell dimension were uncorrelated (Pearson's $r = 0.016$, $t = 0.1652$, df = 114, $p = 0.869$). The effect of shell size on the sampling intensity of species and the interaction of rarity and shell size were significant when outliers (*Spelaeodiscus triarius*, *Pagodulina pagodula*, *Pomatias rivularis*, *Faustina illyrica*, and *Kovacsia kovacsi*) were included in the model. In this model, intermediate sized and rare species (all species listed above except for *F. illyrica*) were more intensively collected than other similar species. The effect of rarity and shell size were significant but the interaction was not significant when outliers were removed (Table 1). According to pairwise comparisons, sampling intensity of the species smaller than 3 mm (species of *Carychium*, *Vallonia*, *Truncatellina*, *Vertigo* and *Vitrea*, and *Punctum pygmaeum*, *Acanthinula aculeata*, *Platyla polita*, *Pupilla triplicata*, *Columella edentula*) was significantly ($p < 0.05$) lower than that of the species bigger than 15 mm (Fig. 5a). Sampling intensity of the rarest fifth of the species was significantly ($p < 0.05$) lower than that of the rest of the species (Fig. 5b).

Discussion

Regional differences in the richness corrected sampling intensity were statistically significant, but regions were not a good predictor of sampling intensity. The Lesser Plain was over-collected relative to its average species richness, probably due to the intensive monitoring

of the Szigetköz area. The spatial pattern of species richness showed sub-regional clustering according to the high autocorrelation within 45-60 km distance. Richness corrected sampling intensity and the residuals after accounting for regional differences were autocorrelated in a much shorter (15 km) distance. These indicate that sampling intensity is heterogeneous within regions and even sub-regions (i.e., within certain hills or mountains).

Areas of high sampling intensity were located around towns and in mountain regions. Some of the towns have particular institutions (museums, universities) or people with malacological interests (Pintér 1981, 1985), for example, Békéscsaba where Gy. Kovács worked and Keszthely where I. Pintér worked. Mountains are often located near large towns (e.g., Budapest and Pécs) and are favourite collecting targets of malacologists (e.g., the Bükk Mountains). The relatively low sampling intensity in the lowlands reflects the predominance of agricultural areas, which are less attractive to malacologists than species rich areas with more natural vegetation.

However, the higher sampling intensity around cities is a more general pattern than just the effect of certain towns or collectors, probably reflecting easier accessibility of areas close to towns where infrastructure is more developed (Reddy and Dávalos 2003). In some cases (e.g., Budapest and Pécs), geomorphological complexity may also contribute to higher sampling intensity (Kühn et al. 2004).

The malacological attratctiveness of protected areas relative to unprotected ones might be a reason for higher sampling intensity within protected areas. Research in national parks also contributed to increased sampling intensity in protected areas. Some species were characterised by high sampling intensity outside protected areas with >25% of the occurrences of *Platyla polita*, *Truncatellina callicratis*, *Macrogastra plicatula* and *Bulgarica vetusta* in unprotected areas. This relatively high proportion suggests that these species are likely to be missed by habitat protection. More than 33% of the occurrences of three species (*Urticicola umbrosus*, *Kovacsia kovacsi*, *Helix lutescens*) were in unprotected areas. The proportion of the occurrences outside protected areas (protection-by-reserves scores, Sólymos 2007) can be used to prioritise species that are missed by habitat protection. Because a high proportion of occurrences of common species are often outside of reserves, the best practice is to incorporate rarity and protection-by-reserves values into an additive index (Sólymos 2007), or to use the indices sequentially as filters to pick up species that meet rarity criteria and are not protected by existing reserves.

The bias due to spatial preference can be reduced *a priori* through well designed and representative sampling of the entire study area by sampling previously underrepresented areas or habitat types, i.e. degraded habitats and agricultural areas. The comparison of degraded and natural systems is also a prerequisite to the modelling the impacts of future land use changes on individual species and biodiversity (Bomhard et al. 2005).

Bias can also be reduced by correcting the data *a posteriori*. Regions with poor data may cause problems in systematic conservation planning when the principle of complementarity is applied. However, correction cannot be complete although to some degree it is necessary to avoid artefacts, i.e., reserve network variability (Freitag and Van Jaarsveld 1998) and estimated spatial costs (Reddy and Dávalos 2003). Species richness within spatial units can be corrected *a posteriori* by using smoothed surfaces (e.g., Williams et al. 1996), rarefaction curves (e.g., Prendergast et al. 1993) or fitting collector's curves to the data (e.g., Soberón and Llorente 1993).

Species richness alone is not a universal guide, since area selection methods based on complementarity also require the identity of the species. Problems may arise when our aim is to model or predict distributions of rare species (or even of undescribed species; see e.g., Bini et al. 2006) because occurrences of rare and dispersal limited species cannot be precisely predicted within potentially suitable habitats (Ponder et al. 2001).

Some rare species are over-represented in the museum records relative to others, but for the majority of species representation in the records is related to the rarity and commonness with common species having more records.

Shell size has a significant effect on the sampling intensity of species. Sampling intensity of microsnails (<3 mm), which can be collected effectively only from leaf litter and soil samples, was significantly lower than that of the more visible species in the largest shell dimension class (>15 mm), which can be collected effectively by direct search. Different selectivity (systematic errors) of soil sampling and direct search is responsible for the observed differences (Cameron and Pokryszko 2005, Sólymos et al. 2007). However, some of the smallest species showed extremely high sampling intensity because of their faunistic importance and rarity.

Rarity and shell dimension might be related in theory negatively because small snails may be more easily transported passively (by wind or by rafting) than large ones (Vágvölgyi 1975, Kirchner et al. 1997). Consequently, small snails should be more widespread. In this case, however, these variables were uncorrelated.

The occurrence of dead and subfossil shells in collections is a special source of bias in the study of land snails (Rundell and Cowie 2003). For example, subfossil shells of *Vallonia enniensis* are common in collections. Treating these shells as an indicator of living populations can be misleading and causes overestimation of the real area of occupancy (non-apparent rarity; Gaston 1994).

Slugs were not analysed in this paper, but are usually underrepresented in museum collections because of collecting and identification problems. Their sampling intensity cannot be compared directly to that of the testaceous snails. The collection of soil dwelling species that burrow into the soil (e.g., semislugs in the genera *Vitrina*, *Semilimax*, *Daudebardia* with fragile shells, and the soil dwelling *Ceciliodes* species and *Mediterranea hydatina*) can also be inefficient without soil sampling, and this might cause pseudo-rarity (Gaston 1994).

The sources of taxonomic bias can be reduced *a priori* by sampling of underrepresented target taxa (i.e. soil dwelling species, semislugs) or *a posteriori* by modelling species distributions (e.g., Ponder et al. 2001) or by using correction factors in the assessment of the rarity of the species (e.g., Sólymos and Fehér 2005, Fehér et al. 2006).

Sampling intensity varied according to regions and possess also small scale (15 km) variation, and it varied according to attributes of the collected taxa, i.e. shell size and rarity. In other words, some areas (accessible and attractive ones) and taxa (common and large ones) were more intensively collected than others. Overcollection might be due to preferences of collectors (including many amateur malacologists) and contracting parties according to the higher sampling intensity around towns and within reserves, respectively. The collecting process would be more cost effective with the same amount of collection data but with more even sampling intensity across space and taxa. Thus, geographic and taxonomic bias in distribution data might be considered as a "biotic impediment". Without planning, these trends in data collecting will probably continue in the furure, increasing the impediment. A detailed understanding of causes of geographic and taxonomic bias may help to properly correct available data and to guide further inventories.

Acknowledgements: Thanks are due to Z. Varga, P. Sümegi, Z. Barta, R. A. D. Cameron and two anonymous reviewers for valuable comments, Z. Fehér for the use of Mollusca database of the Hungarian Natural History Museum and T. Magura for the help in identifying the location of the protected areas. The research was supported by the National R&D Programme ("The origin and genesis of the fauna of the Carpathian Basin: diversity, biogeographic hotspots and nature conservation significance"; Contract no: 3B023-04). During the preparation of the manuscript, P.S. was a postdoctoral fellow of the Hungarian Ministry of Education.

References

Bank, R. 2004. Fauna europaea: mollusca, gastropoda. *Fauna Europaea version 1.2*, http://www.faunaeur.org. Accessed: January, 2006.

Benjamini, Y., D.Drai, G. Elmer, N. Kafkafi, and I. Golani. 2001. Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research* 125:279-284.

Bini, L.M., J.A.F. Diniz-Filho, T.F.L.V.B. Rangel, R.P. Bastos and M.P. Pinto. 2006. Challenging Wallacean and Llinnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. *Diversity and Distributions* 12:475-482.

Bivand, R. (with contributions by many authors). 2007. spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.4-9. URL: http://cran.r-project.org.

Bomhard, B., D.M. Richardson, J.S. Donaldson, G.O. Hughes, G.F. Midgley, D.C. Raimondo, A.G. Rebelo, M. Rouget and W. Thuiller. 2005. Potential impacts of future land use and climate change on the red list status of the Proteaceae in the Cape floristic region, South Africa. *Global Change Biology* 11:1452-1468.

Brooks, T.M., G.A.B. Da Fonseca and A.S.L. Rodrigues. 2004. Protected areas and species. *Conservation Biology* 18:616-618.

Brown, J.H. and M.V. Lomolino. 1998. *Biogeography*. Sinauer Press, Sunderland, Massachusetts.

Cameron, R.A.D. and B.M. Pokryszko. 2005. Estimating the species richness and composition of land mollusc communities: problems, consequences and practical advice. *Journal of Conchology* 38:529-547.

Cowling, R.M., A.T. Knight, D.F. Faith, S. Ferrier, A.T. Lombard, A. Driver, M. Rouget, K. Maze and P.G. Desmet. 2004. Nature conservation requires more than a passion for species. *Conservation Biology* 18:1674-1676.

Dévai, G. and M. Miskolczi. 1987. Javaslat egy új környezetminősítő értékelési eljárásra a szitakötők hálótérképek szerinti előfordulási adatai alapján. (Proposal for a new method of environmental quality evaluation on the basis of grid maps of distribution data of dragonflies). *Acta Biologica Debrecina* 19:33-54.

Fehér, Z. and A. Gubányi. 2001. *The distribution of Hungarian molluscs. The Catalogue of the Mollusca Collection of the Hungarian Natural History Museum*. Hungarian Natural History Museum, Budapest.

Fehér, Z., G. Majoros and A. Varga. 2006. A scoring method for the assessment of rarity and conservation value of the aquatic molluscs in Hungary. *Heldia* 6:101-114.

Freitag, S. and A.S. Van Jaarsveld. 1998. Sensitivity of selection procedures for priority conservation areas to survey extent, survey intensity and taxonomic knowledge. *Proceedings of the Royal Society B* 265:1475-1482.

Gaston, K.J. 1994. *Rarity*. Chapman and Hall, London.

Hungarian Central Statistical Office (KSH) 2001. A Magyar Köztársaság helységnévtára. Központi Statisztikai Hivatal, URL: http://helynevtar.ksh.hu/

Kerney, M.P., R.A.D. Cameron and J.H. Jungbluth. 1983. *Die Landschnecken Nord- und Mitteleuropas*. P. Parey, Hamburg-Berlin.

Kirchner, C., R. Krätzner and F.W. Welter-Schultes. 1997. Flying snails - how far can *Truncatellina* (Pulmonata: Vertiginidae) be blown over the sea? *Journal of Molluscan Studies* 63:479-487.

Kühn, I., R. Brandl and S. Klotz. 2004. The flora of German cities is naturally species rich. *Evolutionary Ecology Research* 6:749-764.

Lomolino, M.V. 2004. Conservation biogeography. In: M.V. Lomolino and L.R. Heaney (eds), *Frontiers of Biogeography: New Directions in the Geography of Nature*. Sinauer Associates, Sunderland, Massachusetts. pp. 293-296.

Miskolczi, M., G. Dévai, G. Kertész and Á. Bajza. 1997. A magyarországi helységek kódjegyzéke az UTM rendszerű 10×10 km beosztású hálótérkép szerint. (Coding list of Hungarian settlements according to a 10×10 km scale grid map of the UTM system.) *Acta Biologica Debrecina* 8:43-194.

Nelson, B.W., C.A.C. Ferreira, M.F.D. Silva and M.L. Kawasaki. 1990. Endemism centres, refugia and botanical collection density in Brasilian Amazonia. *Nature* 345:714-716.

Patten, M.A. and R.A. Erickson. 2001. Conservation value and rankings of exotic species. *Conservation Biology* 15:817-818.

Pintér, I. 1981. A malakológiai felkutatottság mértékének kiszámítása becsléssel. (Estimating the status of malacological inventories.) *Soosiana* 9:29-32.

Pintér, I. 1985. Tájékoztató a magyarországi recens puhatestűek kutatásának eddigi eredményéről. (Information on the results of studying recent molluscs of Hungary.) *Malacological Newsletter* 5:23-28.

Pintér, L. and A.S. Szigethy. 1979. Die Verbereitung der rezenten Mollusken Ungarns: neunachweise und berichtigungen, I. *Soosiana* 7:97-108.

Pintér, L. and A.S. Szigethy. 1980. Die Verbereitung der rezenten Mollusken Ungarns: neunachweise und berichtigungen, II. *Soosiana* 8:65-80.

Pintér, L., A. Richnovszky and A.S. Szigethy. 1979. D*istribution of the Recent Mollusca of Hungary*. Budapest. *Soosiana Suppl. 1*

Ponder, W.F., G.A. Carter, P. Flemons and R.R. Chapman, 2001. Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology* 15:648-657.

Potvin, C. and D.A. Roff. 1993. Distribution-free and robust statistical methods: viable alternatives to parametric statistics. *Ecology* 74:1617-1628.

Prendergast, J.R., S.N. Wood, J.H. Lawton and B.C. Eversham. 1993. Correcting for variation in recording effort in analyses of diversity hotspots. *Biodiversity Letters* 1:39-53.

Pressey, R.L. 2004. Conservation planning and biodiversity: assembling the best data for the job. *Conservation Biology* 18:1677-1681.

R Development Core Team. 2007. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

Ramsey, R.D. and L. Shultz. 2004. Evaluating the geographical distribution of plants in Utah from the atlas of vascular plants of Utah. *Western North American Naturalist* 64:421-432.

Reddy, S. and L.M. Dávalos. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography* 30:1719-1727.

Rundell, R.J. and R.H. Cowie. 2003. Preservation of species diversity and abundances in Pacific island land snail death assemblages. *Journal of Conchology* 38:155-170.

Soberón, J.M. and J.B. Llorente. 1993. The use of species accumulation functions for the prediction of species richness. *Conservation Biology* 7:480-488.

Sokal, R.R. and F.J. Rohlf. 1995. *Biometry*. Freeman and Co, New York.

Sólymos, P. 2007. Are current protections of land snails in Hungary relevant to conservation? *Biodiversity and Conservation* 16:347-356.

Sólymos, P. and Z. Fehér. 2005. Conservation prioritization using land snail distribution data in Hungary. *Conservation Biology* 19:1084-1094.

Sólymos, P., I. Czentye and B. Tutkovics. 2007. A comparison of soil sampling and direct search in malacological field inventories. In: K. Tajovský, J. Schlaghamerský and V. Pi□l (eds), *Contributions to Soil Zoology in Central Europe, II*. České Budéjovice. pp. 161-163.

Stohlgren, T.J., J. F. Quinn and G. S. Waggoner. 1995. Status of biotic inventories in US national parks. *Biological Conservation* 71:97-106.

Vágvölgyi, J. 1975. Body size, aerial dispersal, and origin of the Pacific land snail fauna. *Systematic Zoology* 24:465-488.

Whittaker, R.J., M.B. Araújo, P. Jepson, R.J. Ladle, J.E.M. Watson and K.J. Willis. 2005. Conservation biogeography: assessment and prospect. *Diversity and Distributions* 11:3-23.

Williams, P., D. Gibbons, C. Margules, A. Rebelo, C. Humphries and R. Pressey. 1996. A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology* 10:155-174.

Williams, P.H., C.R. Margules and D.W. Hilbert. 2002. Data requirements and data sources for biodiversity priority area selection. *Journal of Biosciences* 27:327-338.

Received November 18, 2007
Revised November 29, 2007
Accepted December 1, 2007